ORIGINAL PAPER

Educational institutions and the integration of migrants

Nicole Schneeweis

Abstract In this paper, I study educational integration of students with migration background using data from five international student assessment studies. First, Blinder–Oaxaca decompositions are used to allow for a comparison of integration of migrant students across countries and time. In a second step, integration is related to institutional characteristics of the schooling system. Pooled, country-group and country fixed effects estimations show that time in school and early education are positively related to the integration of students with migration background. Furthermore, in the OECD countries, educational integration in science is positively related to external student assessment policies.

1 Introduction

Educational integration is an important precondition for the economic assimilation of immigrants in the host societies. International student assessment studies cause concern about the integration of immigrant children in schools. The Program for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS) show that students

N. Schneeweis (✉)
Department of Economics, Johannes Kepler University,
Aubrunnerweg 3a, 4040 Linz, Austria
e-mail: nicole.schneeweis@jku.at

⌀ Springer

who were born abroad perform significantly worse in the achievement tests, compared to native students. The average achievement gaps in the OECD-countries amount to about 25 test score points in math and 28 in science (25% and 28% of the standard deviations in test scores).

Several studies investigate the achievement gaps in more detail. Entorf and Minoiu (2005) have shown that not only the PISA achievement gaps between migrants and non-migrants vary substantially across OECD countries, but also the socioeconomic background of the immigrants and its influence on achievement. Ammermüller (2007) has raised the question of why immigrants in Germany performed so poorly in PISA. The answer is twofold: immigrants in Germany come from less favorable social backgrounds and they get lower returns to their characteristics than German natives.

Why do we observe large gaps in cognitive skills between students with foreign background and native students? Can these gaps be explained by differences in student characteristics? And most importantly, what can policy do? How should schooling be organized to further the integration of children with foreign backgrounds?

This essay is aimed at quantifying the disadvantage of immigrant children in education and relating it to institutional conditions of the education system. In the first step, educational integration of immigrants and second-generation immigrants is measured and made comparable across countries and time, using micro-data of several international student assessment studies. In the second step, I estimate the effects of certain characteristics of the education system, such as pre-primary education, time in school, or the segregation of students among schools, on the mean level of educational integration based on a cross-country time-series analysis.

2 The integration of immigrants

The raw data of various achievement tests give a substantial drawback for students with migration backgrounds. These achievement gaps cannot be compared directly across countries, since educational success is largely determined by the social background of the students (Hanushek and Luque 2003; Wößmann 2005a) and different countries have different immigrant populations. Depending on the income situation, the geographic region, the immigration policy, and many other characteristics, they attract migrants with different abilities and social backgrounds.

I use the Blinder–Oaxaca decomposition to construct a measure of integration that is comparable across countries and time (Blinder 1973; Oaxaca 1973). The mean achievement gap between natives and migrants in a country is decomposed into a part that is explained by differences in social background characteristics and a part that remains unexplained. Educational production functions are estimated separately for natives, immigrants, and

second-generation immigrants. The average native, immigrant, and second-generation immigrant test scores ($\overline{Y}_n, \overline{Y}_i, \overline{Y}_s$) can be written as products of the estimated coefficients, including the intercepts ($\hat{\beta}_n, \hat{\beta}_i, \hat{\beta}_s$) and the average endowments ($\overline{X}_n, \overline{X}_i, \overline{X}_s$) of the three groups:

$$\overline{Y}_n = \hat{\beta}_n \overline{X}_n, \quad \overline{Y}_i = \hat{\beta}_i \overline{X}_i \text{ and } \overline{Y}_s = \hat{\beta}_s \overline{X}_s. \tag{1}$$

The average achievement gaps between students with foreign background and native students can be formulated as

$$\Delta \overline{Y}_{i-n} = \hat{\beta}_i \overline{X}_i - \hat{\beta}_n \overline{X}_n = \underbrace{\hat{\beta}_n \left(\overline{X}_i - \overline{X}_n \right)}_{explained} + \underbrace{\overline{X}_i \left(\hat{\beta}_i - \hat{\beta}_n \right)}_{unexplained}. \tag{2}$$

$$\Delta \overline{Y}_{s-n} = \hat{\beta}_s \overline{X}_s - \hat{\beta}_n \overline{X}_n = \underbrace{\hat{\beta}_n \left(\overline{X}_s - \overline{X}_n \right)}_{explained} + \underbrace{\overline{X}_s \left(\hat{\beta}_s - \hat{\beta}_n \right)}_{unexplained}. \tag{3}$$

The explained part of the test score gap considers that students with foreign background may be endowed with less favorable socioeconomic characteristics and, therefore, may be less successful in education. The unexplained part of the achievement gap can be interpreted as a measure of integration. Multiplying the unexplained by -1 gives the answer to the following question: By how many test score points would immigrants perform better, given their own endowments, if they had the same returns as native students?[1] As mentioned above, a similar approach was used by Ammermüller (2007) to study the PISA achievement gap between migrant students and native Germans.

The separate estimation of the achievement function for the three groups allows heterogenous returns to individual characteristics for natives and students with migration background. It is plausible to assume that natives and foreign students are different populations and obtain different returns to their endowments. A high educational attainment of parents, for example, might not have the same positive impact for migrant students as for natives. Similarly, studies on the returns to education on the labor market show that individuals with migration background get a significantly smaller payoff to their education (Chiswick and Miller 2008; Hartog and Zorlu 2009).

The unexplained is interesting to analyze and compare across countries; nevertheless, it is problematic for the purpose of this paper. The analysis of institutional effects needs a measure of integration that is comparable across countries and does not depend on the average characteristics of immigrant students in a certain country. The question the measure should be able to

---

[1]Priority is given to $\hat{\beta}_n$ and $\overline{X}_{i/s}$ over $\hat{\beta}_{i/s}$ and $\overline{X}_n$ as weights. It is assumed that migrants face drawbacks in education rather than natives are favored (and would get lower returns in the absence of migrant students).

answer is the following: *How much better would a representative student with foreign background perform in a given institutional regime if he or she had the same returns as the native students in that regime?*

The unexplained test score gaps are, therefore, standardized:

$$I_i = \overline{\mathbf{X}}_{\mathbf{i}}^{\mathbf{st}}(\hat{\beta}_{\mathbf{i}} - \hat{\beta}_{\mathbf{n}}) \text{ and } I_s = \overline{\mathbf{X}}_{\mathbf{s}}^{\mathbf{st}}(\hat{\beta}_{\mathbf{s}} - \hat{\beta}_{\mathbf{n}}), \tag{4}$$

where $\overline{\mathbf{X}}_{\mathbf{i}}^{\mathbf{st}}$ and $\overline{\mathbf{X}}_{\mathbf{s}}^{\mathbf{st}}$ are the mean characteristics of immigrants and second-generation immigrants in the whole sample. The sensitivity of the results on institutional effects to this standardization is discussed later on.

Note that the measure of integration is a relative one. It gives the drawback of students with migration backgrounds, relative to the native students in that country. This is exactly the measure I need to represent the situation of immigrants. It is not important whether immigrants in the USA do worse than German natives or the average native in the sample. The important question is the relative position of immigrant students in the societies of their host countries, where they are going to live and work.

I include individual and family background variables in the achievement regressions. School resources, like class size or teacher characteristics, may play an important role but are not randomly allocated across schools, just as little as students with migration background are. The allocation of school resources is a source of integration policy and controlling for school characteristics in the educational production functions would overestimate the true level of integration.

2.1 Data from PISA and TIMSS

I use micro-data from several waves of two different international student assessment studies. TIMSS has been conducted by the International Association for the Evaluation of Educational Achievement (IEA) in 1995, 1999, and 2003 in about 50 different countries and PISA has been organized by the OECD in 2000 and 2003. In both surveys, about 4,000 secondary education students from about 170 schools were assessed in each participating country in each wave. Among other things, the surveys provide estimates of student proficiencies in mathematics and science, as well as detailed background information of students and schools.

My sample consists of 167 country-years, which span a time period of 9 years (from 1994 to 2003). See Table 1 for a list of the countries. Some country-years were dropped from the sample because of missing background information. Since the decomposition approach is based on separate estimations for immigrants, second-generation immigrants, and natives in each country, I further dropped all country-years with less than 40 students in a group. As will be explained below, I calculated the standard errors of the decompositions and use the inverse as weight for the regressions in the second part of the paper to account for differences in the number of observations.

Table 1 List of countries used in the analysis

Country	ISO-Code	Study-Years
Southern European countries		
Spain	ESP	t1995, p2000, p2003
Greece	GRC	t1995, p2000, p2003
Italy	ITA	t1995, t1999, p2000, t2003, p2003
Portugal	PRT	t1995, p2000, p2003
Middle and Northern European countries		
Austria	AUT	t1995, p2000, p2003
Belgium flemish	BFL	t1995, t1999, p2000, t2003, p2003
Belgium french	BFR	t1995, p2000, p2003
Switzerland	CHE	t1995, p2000, p2003
Germany	GER	t1995, p2000, p2003
Denmark	DNK	t1995, p2000, p2003
Finland	FIN	p2000, p2003
France	FRA	p2000, p2003
Iceland	ISL	t1995, p2000, p2003
Luxembourg	LUX	p2000, p2003
Netherlands	NLD	t1995, t1999, p2000, t2003, p2003
Sweden	SWE	t1995, p2000, t2003, p2003
Eastern European countries (relatively rich)		
Czech Republic	CZE	t1995, t1999, p2003
Estonia	EST	t2003
Hungary	HUN	t1995, t1999, t2003, p2003
Slovak Republic	SVK	t1995, t1999, t2003, p2003
Slovenia	SVN	t1995, t1999, t2003
Eastern Europe and Russia		
Bulgaria	BGR	t1999, t2003
Lithuania	LTU	t1995, t2003
Latvia	LVA	t1995, t1999, p2000, t2003, p2003
Macedonia	MKD	t1999, t2003
Moldova	MDA	t1999, t2003
Romania	ROM	t1995
Russia	RUS	t1995, t1999, p2000, t2003, p2003
Serbia	YUG	t2003, p2003
English speaking countries		
Australia	AUS	t1994, t1998, p2000, t2002, p2003
Canada	CAN	t1995, t1999, p2000, p2003
England	ENG	p2000, t2003, p2003
Ireland	IRL	t1995, p2000, p2003
New Zealand	NZL	t1994, t1998, p2000, t2002, p2003
Scotland	SCO	t1995, t2003, p2003
United States	USA	t1995, t1999, p2000, t2003, p2003
South America and Mexico		
Chile	CHL	t1998, t2002
Colombia	COL	t1995
Mexico	MEX	p2000, p2003
Uruguay	URY	p2003
Near East		
Armenia	ARM	t2003
Bahrain	BHR	t2003
Cyprus	CYP	t1995, t1999, t2003
Iran	IRN	t1995, t2003
Israel	ISR	t1995, t1999, p2002, t2003
Jordan	JOR	t1999, t2003
Kuwait	KWT	t1995
Lebanon	LBN	t2003

Table 1 (continued)

Country	ISO-Code	Study-Years
Saudi Arabia	SAU	t2003
Turkey	TUR	t1999, p2003
Far East		
Hong Kong	HKG	t1995, t1999, p2002, t2003, p2003
Indonesia	IDN	t2003
Macao, China	MAC	p2003
Malaysia	MYS	t1998, t2002
Philippines	PHL	t1995, t1999, t2003
Singapore	SGP	t1994, t1998, t2002
Thailand	THA	t1995, t1999
Africa		
Egypt	EGY	t2003
Ghana	GHA	t2003
Morocco	MAR	t1999, t2003
South Africa	ZAF	t1995, t1998, t2002
Tunisia	TUN	t1999, p2003
# Countries = 62 (OECD = 28)		
# Country − years = 167 (OECD = 95)		
# Years per country = 2.7 (OECD = 3.4)		

The TIMSS waves 1995, 1999 and 2003 were implemented in the years 1994, 1998, and 2002 in some countries and some countries carried out the PISA 2000 assessment in 2002

For each of the 167 country-years, I estimate integration of immigrants and second-generation immigrants (I_i, I_s). The dependent variable in the underlying educational production function is the student test score in PISA and TIMSS, respectively, and individual student characteristics are *age, grade, sex, the highest education level obtained by parents, the number of books at home, whether students have a computer, a calculator, and a desk to study at home and whether they speak the national language at home.* Table 2 gives summary statistics and a description of the student-level variables.[2]

This rich list of explanatory variables represents the individual characteristics of the students and their family background. Some more variables concerning the immigration status, like the reasons why the families migrated, the number of years since immigration, and the home countries of the immigrants cannot be observed in most data sets. A variable that is seen to play an important role for the economic assimilation of immigrants is whether the students speak the national language at home. This variable is available in the data and included in the achievement regressions. Since this variable is critical and may be correlated with unobserved characteristics of the host country, a sensitivity check without this control is presented in the second part of the paper.

The achievement functions are estimated with survey regressions, with students weighted according to sampling probabilities, and the dependence

---

[2]For some students, not all explanatory variables are available. Since migration status, education of parents, and grade at school are important, students with missing values are dropped. For all other variables, missing dummies are included in the educational production functions. For these variables, the number of missings is small and ranges from 0.06% to 3.37%.

Table 2 Student-level variables and decomposition results

Variable	Description	Mean	Stdev
Test score			
Math score	Transformed plausible value of math proficiency	500	100
Science score	Transformed plausible value of science proficiency (science sample)	500	100
Ethnicity			
Immigrant	Student was born in a foreign country	0.082	
Second-generation immigrant	Student's father, mother or both were born in a foreign country and student was born in the country	0.132	
Native	Student and his/her parents were born in the country	0.786	
Individual characteristics			
Age	Age of student in years	14.779	1.050
Grade	Grade at school	8.480	1.046
Female	Student is female	0.509	
Number of books at home			
Books1	None–10 books	0.123	
Books2	11–25 books (1–50 in PISA 2000)	0.201	
Books3	26–100 books (51–100 in PISA 2000)	0.276	
Books4	101–200 books (101–250 in PISA 2000)	0.177	
Books5	More than 200 books (more than 250 in PISA 2000)	0.222	
Highest education level reached by a parent			
Isced01	No schooling or primary education	0.122	
Isced2	Lower secondary education	0.099	
Isced34	Upper secondary education	0.491	
Isced56	Tertiary education	0.287	
Computer	Student has a computer at home	0.608	
Calculator	Student has a calculator at home	0.935	
Study desk	Student has an own desk to study at home	0.876	
National language	Student speaks the test language, another national language or a national dialect at home (PISA), student speaks the test language at home (TIMSS)	0.853	
Decomposition results			
$I_{i,s}$ math	Integration in math	−5.838	32.293
$I_{i,s}$ science	Integration in science	−5.448	36.191
I_i math	Integration of immigrants	−6.651	36.741
I_s math	Integration of second-generation immigrants	−5.025	27.215
I_i science	Integration of immigrants	−8.973	44.277
I_s science	Integration of second-generation immigrants	−1.923	25.337
# Students in the mathematics sample = 753, 282			
# Students in the science sample = 753, 445			
# Country-years (decomposition results) = 167 for immigrants, 167 for second-generation immigrants			

The means and standard deviations of the student-level variables are based on the math sample and weighted according to the students' sampling probabilities. The weights are adjusted to ensure an equal contribution of each country-year

of standard errors within clusters (schools) is taken into account. Part of the difference in the study designs between PISA and TIMSS can thereby be eliminated.

PISA and TIMSS are of similar type; both are aimed at obtaining an internationally comparable measure of the proficiency of secondary school students and both incorporate a comparable quality standard with respect to the design and implementation of the assessment.[3] The similarity of the PISA and TIMSS survey designs allows the use of both studies together. Concentrating on PISA or TIMSS only, is not possible in this study because sample size would be too small, especially for the country-fixed-effects regressions in the second part of the paper. See the Appendix for a short description of differences between PISA and TIMSS, as well as the applied transformation strategy to reach comparability of student achievement scores. A more detailed analysis of the comparability of PISA and TIMSS is given in Brown et al. (2007).

2.2 Integration in various countries

This section summarizes the actual (non-standardized) results of the Blinder–Oaxaca decompositions in mathematics and science. Figure 1 shows the total achievement gaps between foreign students and natives, decomposed into an explained and an unexplained part. Due to the wide range of different countries, these are arranged into nine country groups, wherefrom mean values are reported.

On average, students with foreign backgrounds achieve lower scores than native students, and a part of the test score gap can be explained with differences in student characteristics in each country group. The total gaps are similar in math and science and range from about −34 points in Middle and Northern Europe to about −2 in the Near East. Remember, the test scores are normally distributed with a (weighted) mean of 500 and a (weighted) standard deviation of 100 in math and science.

The Middle and Northern European countries show large achievement gaps, −33 in math and −35 in science, with about 15% and 26% remaining unexplained. In Southern Europe, both gaps are about −10, and very small fractions remain unexplained. In Eastern Europe, the gaps are higher in relatively rich countries and the unexplained is positive, meaning that foreign students receive higher returns to their characteristics than natives. This is the same in the English-speaking countries, which have rather small total gaps.

An interesting pattern arises when comparing the integration measures with and without controlling for the national language proficiency in the achievement regressions (these comparisons are not shown in the figures). In Middle and Northern Europe 36% and 43% and in Southern Europe 32% and 50% of the gaps remain unexplained. Furthermore, in the relatively rich Eastern European countries, as well as in the English-speaking countries, the unexplained differentials turn zero in math and negative in science. Hence, the

---

[3]See (OECD 2001, 2002, 2004, 2005) and IEA (http://TIMSS.bc.edu/isc/publications.html) for detailed information on the PISA and TIMSS surveys.
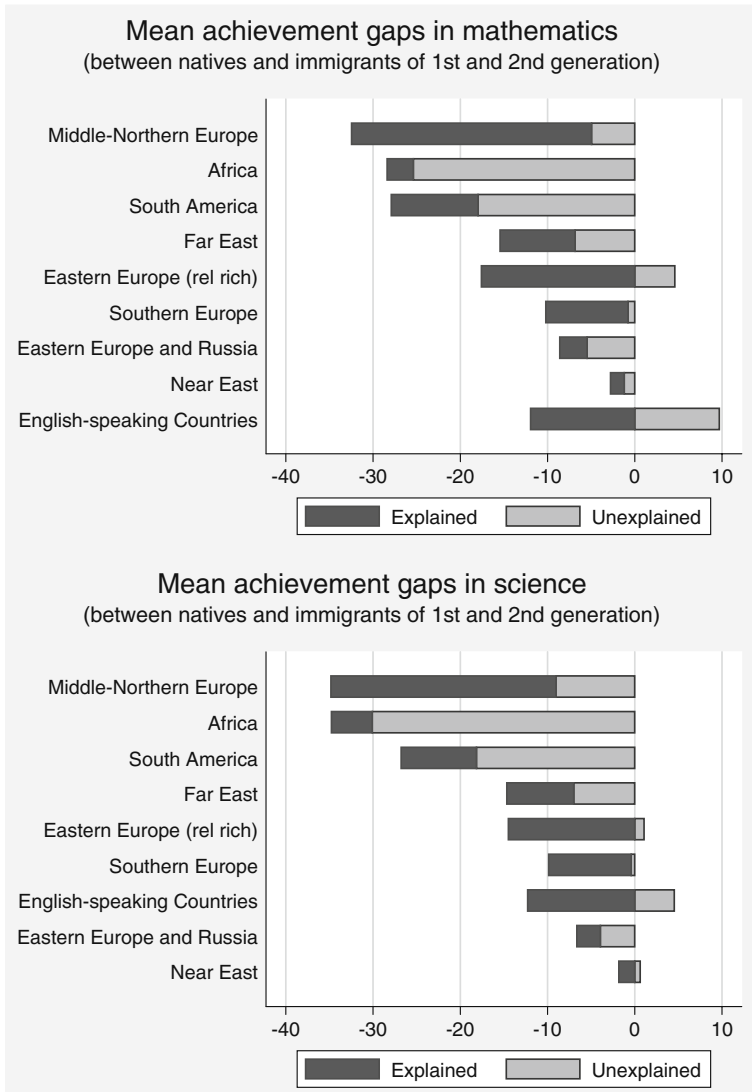
Fig. 1 Achievement gaps by country groups

proficiency of the national language is a major vehicle for migrant students to catch up in education. This result is an important finding, since language proficiency can be influenced by public policy in different ways, like the provision of special language courses in schools or language trainings for adult immigrants.

Most countries in the sample are members of the OECD and the decomposition results for these countries are shown separately in Fig. 2. In addition
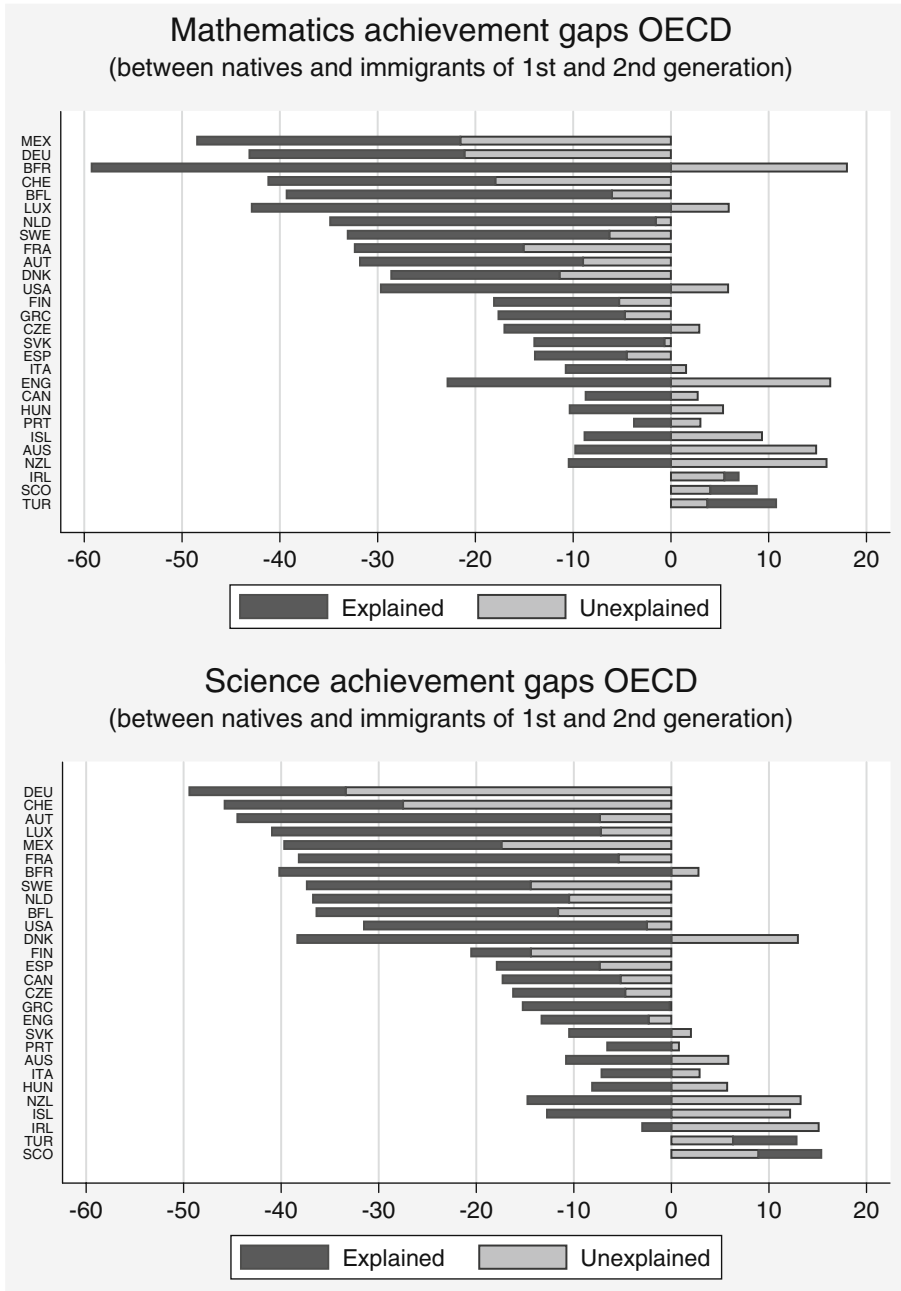
Fig. 2 Achievement gaps in math and science in the OECD

to the large variation in achievement gaps and unexplained differentials, the graphs tell us two important stories:

- *The English-speaking countries* are found in the middle and lower tail of the gap distribution. In these countries, the unexplained test score gap is always positive in math and mostly positive in science; thus, children with migration background get higher returns to their characteristics than native students. One explanation for this result might be the fact that English is a world language and, therefore, it might be easier for migrants to integrate in such countries. Furthermore, Australia, Canada, New Zealand, and the USA are frequently characterized as traditional countries of immigration, and most of these countries follow a selected immigration policy, targeted at highly educated individuals with professional skills and adequate language proficiency (Miller 1999; Entorf and Minoiu 2005).
- *Within Continental Europe*, the German-speaking countries, the Benelux countries, France, and the Scandinavian countries can be found in the upper part of the distribution and the Southern and Eastern European countries are ranked in the lower tail.

As mentioned above, a standardized version of the unexplained differential is used as measure of integration in the regression analysis of the second part. Table 2 gives summary statistics of the standardized unexplained test score gaps in math and science. In both subjects, second-generation immigrants face a higher level of integration than first-generation immigrants. This result was expected as the assimilation of immigrants is associated with their length of stay in the host country.

The unexplained part of the test score gap can be interpreted as a measure of integration, since, multiplied by -1, it tells us how much better students with migration backgrounds would perform if they had the same returns as native students. The measure is not reliable and cannot be compared if unobserved ability differences between natives and immigrants in the various countries exist that influence the test scores of the students.

2.3 Unobserved ability differences?

The immigrants of a given country are a highly selected group of people. Certain factors motivated their decision to migrate, while others decided to stay in their country. Economic models have been developed that investigate the selectivity of *economic migrants* with respect to their ability. The most important is the Roy model, applied by Borjas (1987, 1999) and extended by Chiswick (1999). This migration model assumes that the rate of return from migration is different for high-ability and low-ability individuals and determines whether an individual decides to migrate. Positive fixed costs of migration lead to a positive selection of migrants, which is intensified if high-ability individuals are more efficient in the migration process. Furthermore, economic immigrants are negatively selected if the wages in the destination,

relative to the home country, are higher for low-ability individuals. This result implies that, for a constant ability distribution across countries, a lower relative income inequality in the destination country negatively selects migrants. In total, due to the costs of migration and the likelihood that such costs are lower for high-ability individuals, economic migrants are positively self-selected. This selectivity is diminished if the relative income inequality is higher in the home country.

Economic reasons are not the only reasons why people migrate. *Refugees* have to move because their safety or freedom is at risk, and other people move to accompany *family members* in other countries. Such migrants are mostly not favorably selected, as studies on unemployment and earnings show (Chiswick 1999). Furthermore, not only does the supply of immigrants determine the foreign population in a country, but demand-side effects are relevant, too. Some countries follow an *immigration policy* that is restricted to well-educated immigrants with good language skills.

Overall, as long as ability and motivation cannot be observed entirely, integration is likely to be over- or underestimated depending on unobserved ability differences. Economic theories predict that, in countries with a relatively low level of income inequality and a big part of immigration due to non-economic reasons, immigrants are likely to be negatively self-selected with respect to their ability. On the contrary, a selective immigration policy leads to a positive selection of immigrants. Thus, the low level of integration in the European countries may be underestimated, while the high (or even positive) level of integration in traditional immigration countries may be overestimated.

3 The role of institutions

What is the influence of the education system and what can policy do to further the integration of students with migration background? To find answers to these questions, integration is related to institutional characteristics of the education system.

Segregation   School systems differ with respect to the segregation of migrants and poor students among schools. A high degree of migrant or social segregation is caused either by selectivity mechanisms of the education system, like general tracking, or by a high degree of residential segregation in comprehensive education systems. On the one hand, immigrants may profit from segregated schools because teachers may be more able to target the needs of the students in more homogenous classes. On the other hand, a higher degree of segregation can harm immigrant children because they have a higher probability of being allocated to low-grade school types and schools (e.g., Rees et al. 1996; Epple et al. 2002). Attending a lower-grade school type or a school in a poor neighborhood can have negative effects for mainly two reasons: school resources might not be equally allocated to schools and the absence of clever classmates and students from supporting homes with caring

parents may have negative effects on the learning climate. Empirical studies on peer effects show that especially low-ability students and students from less favorable family backgrounds profit from being placed in schools and classes with high-ability students or kids from parents with favorable socioeconomic characteristics (Winston and Zimmerman 2003; Sacerdote 2001; Schneeweis and Winter-Ebmer 2007).

Time in school If foreign students spend more time in school, pedagogically supported, together with kids of other ethnic groups, they should interact with each other and learn the national language and other national habits, and integration can take place. A full-time school system should, therefore, lead to a higher degree of integration. On the other hand, especially for students with learning or language problems, too much time in school might be too demanding. Aksoy and Link (2000) investigated US panel data and found mathematics achievement to be positively affected by the number of minutes per math class. The number of legal days in school and hours of school per week show no consistent effects. Lewis and Seidman (1994) found large positive effects of the length of the school year in a cross-section analysis.

Pre-primary education Carneiro et al. (2005) have studied labor market discrimination of ethnic minorities in the USA and argue that deficits in cognitive skills of minorities emerge early and widen with schooling. The authors recommend that policy measures to increase the labor market success of minority groups should be applied as early as possible. Early-childhood programs, like kindergartens, day-care centers, and pre-schools, are aimed at preparing children for primary education and providing an equal starting point for all children. Currie (2001) investigated pre-school programs in the USA and found significant benefits for educational attainment and earnings, especially for disadvantaged children. In another study, Currie and Thomas (1999) focused on the impacts of Head Start, a subsidized pre-school program in the USA. The authors show that all children benefit from Head Start, compared to their siblings who did not attend the program, and Head Start closes one quarter of the test score gap between Hispanic and white children. This evidence on pre-primary education suggests that a country should be more effective in decreasing inequality the more children of immigrants and second-generation immigrants attend pre-primary education.

Entry age of schooling Whether it is better to enroll children in school at an earlier or later time has been discussed frequently. Most economic studies in this regard rely on within-country variation in entry age due to month or quarter of birth (e.g. Angrist and Krueger 1992). In this study, variation in the school entry age across countries is investigated. It is hypothesized that enrollment at age 7 is detrimental for immigrant students, since the integration process in school starts later. Attendance at age 5 should operate the other way around and reduce the drawback of migration.

Pupil–teacher ratio The question of whether class size affects student achievement has been studied extensively. Two prominent studies should be mentioned. Krueger (1999) investigated the Tennessee Student–Teacher Achievement Ratio experiment (STAR) and showed that smaller classes in primary education help students, especially those from disadvantaged families. The same result was found by Angrist and Lavy (1999) for Israeli primary-education students. The authors employed a regression discontinuity design, exploiting random variation in class size due to grade enrollment and a maximum class size rule. The pupil–teacher ratio in primary education, thus, is assumed to have a negative impact on the integration of foreign students.

External student assessment Central examinations restrict the latitude of teachers' grading practices, provide information on the relative standing of students and schools, and induce parental and public pressure on students, teachers, and schools. Thus, central student assessments should be positively related to academic achievement. Wößmann (2005b) has shown that central exams exert heterogenous performance effects and reduce the achievement drawback of children with migration background.

Remedial and enrichment courses Immigrant children should profit from school systems that offer special courses in academic subjects for low achieving students. Enrichment activities for gifted students, on the other hand, may increase the unexplained achievement gap if students with migration backgrounds are less frequently promoted in such programs.

3.1 Analysis of institutional effects

I use pooled weighted least squares and fixed-effects regressions to analyze institutional effects on integration. The model can be written as

$$I_{ct} = \alpha_0 + \alpha_1 \mathbf{E_{ct}} + \alpha_2 \mathbf{Y_{ct}} + \alpha_3 \mathbf{C_{ct}} + v_c + u_{ct}, \qquad (5)$$

where c and t index countries and time. The dependent variable I_{ct} is the standardized unexplained test score gap of immigrants and second-generation immigrants, respectively. The vector $\mathbf{E_{ct}}$ represents educational institutions, $\mathbf{Y_{ct}}$ denotes the income situation of the country, and $\mathbf{C_{ct}}$ is a vector of control variables. The error term is split up into a part that is constant within each country v_c and an idiosyncratic part u_{ct}.

A consistent estimation of α_1 requires that $\mathbf{E_{ct}}$ and the error terms are uncorrelated. A country-fixed-effects model is a perfect way to eliminate the country-specific error term v_c, which, among other things, includes the time invariant unobserved ability composition of the immigrant population. This unobserved ability composition may be correlated, for example, with the segregation of migrants among schools or enrollment in pre-primary education. In a country-fixed effects model, the identifying assumption is reduced to the

condition that foreign students observed in 1994 should not differ from those in 2003 in their unobserved characteristics within each country.

The country-fixed-effects model uses the variation within countries only, capturing a time component and a sampling component. Thus, next to changes over a time span of about 10 years, an additional source of variation is available. The dependent variable and most of the institutional variables are extracted from PISA and TIMSS data, as will be explained in the next section. While, in PISA, 15–16-year-old students are sampled, in TIMSS, grade-8 students are assessed. There is additional variation due to the fact that TIMSS and PISA students face different environments in their schooling systems. The students for which integration is calculated at a given year are confronted with a set of environmental characteristics and exactly this set is related to their levels of integration.

A variance decomposition in between and within components shows that, out of the eight institutional variables, the within variance is above 20% for six and above 30% for two variables. The dependent variables show a within variance component of above 40%. Thus, the within variation should be sufficient to estimate a fixed-effects model.

Potential problems of a country-fixed-effects model are that the coefficients of variables that do not change over time cannot be estimated. Furthermore, differencing out country effects may increase attenuation biases because of measurement errors. Measurement errors might arise in this study from an imprecise measurement of integration, the fact that it is calculated from different student assessment studies and the aggregation of institutional characteristics to the country level. Due to these reasons, three methods are used to estimate the model: pooled weighted least squares, a model with country-group dummies as listed in Table 1, and a country-fixed-effects model.

Equation 5 explicitly includes $\mathbf{Y_{ct}}$, the level of income, as well as income inequality in the host country. The income situation should capture the general availability of resources, as well as the potential selectivity of migrants according to the Roy model. $\mathbf{C_{ct}}$ includes the mean national test score as an overall quality measure of the education system, a dummy for second-generation immigrants, wave dummies, country-group dummies, or country dummies, respectively, and country-group-wave dummies in the country-group, as well as the country-fixed-effects models.

As mentioned above, since the dependent variable is constructed, following Silber and Weber (1999) and Card and Krueger (1992), observations are weighted by the inverse of the disturbance to the dependent variable. Standard errors of the decompositions are obtained by bootstrapping, with 200 bootstrap replications employed, and their inverse is used as weight. Country-years, in which integration is estimated with a lower degree of accuracy, are weighted less in the regressions. This is important because, for example, as the number of migrants is small in some countries, the coefficients of the achievement regressions might not always be significant or the

background characteristics of migrants might be measured with error in some countries.

The model is estimated for the whole sample and for the subsample of OECD countries. Though the OECD sample is rather small, it includes countries that show comparable characteristics, and the identifying assumptions are more likely satisfied. Furthermore, to account for changes in the unobserved characteristics of migrants over time that might be correlated with institutional characteristics, I control for the home countries of the foreign population as a robustness check. For a small number of countries, I have aggregate data on the migration regions of the migrant population stock.

3.2 Explanatory variables

The empirical analysis of institutions is based on data from different sources. See Table 3 for a description and summary statistics of the country-level variables.

First, the PISA and TIMSS databases include useful information on schools, where the relevant school variables are aggregated to the country level.[4] One might ask why school data are aggregated to the country level and their effects on immigrant performance are not estimated directly. Exploiting the variation among schools entails the problem of student self-selection. If high-ability immigrants are more likely to choose better schools with clever peers and adequate equipment, the effects of school resources cannot be identified. Aggregation helps to overcome this identification problem to the cost of measurement errors in focusing only on the mean level of resources, regardless of the distribution.

Segregation of students among schools is measured by the Duncan and Duncan (1955) dissimilarity index, recently applied by Burgess and Wilson (2005) and Jenkins et al. (2008). The dissimilarity index of migrant segregation is based on a binary variable that splits the population into two groups: Migrant segregation $= \frac{1}{2} \sum_{s=1}^{S} |\frac{f_s}{F} - \frac{n_s}{N}|$, where f_s and n_s are the numbers of foreign and native students in school s and F and N are the total numbers in the country. The index ranges from 0 to 1 and gives the fraction of students with migration background that has to be moved to other schools to ensure an equal representation of foreign students in each school. Analogously, a social segregation index is calculated, where the two groups represent students with more and less than 25 books at home. The segregation indices differ between PISA and TIMSS. While PISA sampled single students from schools, TIMSS assessed whole classes. Thus, the PISA data refer to school segregation, whereas the TIMSS data measure segregation among classes. Next to segregation, the information on time in school, external student assessment, and

---

[4]Because PISA and TIMSS do not provide representative samples of schools in a country, the aggregation is based on weighted schools, with the weight for a school as simply the sum of all student weights within that school. Since the student sample is representative for the total student population, weighted school aggregates are good proxies for the school population.

Table 3 Country-level variables

Variable	Description	Mean	Stdev	Obs
PISA and TIMSS data				
Migrant segregation	Dissimilarity index of foreign students in schools (immigrants and second-generation immigrants)	0.386	0.103	167
Social segregation	Dissimilarity index of students with less than 25 books at home in schools	0.339	0.074	167
Hours per year	Mean instructional hours per school year divided by 100	9.318	1.707	155
External studass	Fraction of schools that does not have the primary responsibility for student assessment policies	0.143	0.212	151
Remedial courses	Fraction of schools that provides remedial courses in mathematics (academic subjects) for weak students	0.766	0.185	167
Enrichment courses	Fraction of schools that provides enrichment courses in mathematics (academic subjects) for gifted students	0.512	0.281	167
World development indicators 2005				
GDP	GDP per capita (ppp, in constant 2000 international $) divided by 100	180.630	101.593	162
Gini	Gini coefficient	0.355	0.081	154
Pre-primary enrollment	Gross percentage of students who are enrolled in pre-primary education (lagged: time when students were 4/5 years old)	63.712	29.810	157
Pupil–teacher ratio	Number of pupils per teacher in primary education (lagged: time when students were 7/8 years old)	16.738	6.946	141
UNESCO Institute for statistics				
Age primary 5	Entry age of primary education is 5	0.067		165
Age primary 6	Entry age of primary education is 6	0.679		165
Age primary 7	Entry age of primary education is 7	0.254		165
OECD Trends in international migration				
Migration regions: Fraction of foreign population coming from . . .				
Western Europe	Western European countries	19.046		65
Southern Europe	Southern European countries	3.456		65
Eastern Europe	Eastern European countries	14.481		65
North America	North America	8.375		65
South America	South America	1.553		65
Africa	Africa	7.274		65
Asia	Asia	13.500		65
Oceania	Oceania	2.986		65
Other	Other countries	29.330		65

remedial and enrichment courses is taken from PISA and TIMSS data. Further data sources are the World Banks' World Development Indicators 2005, the UNESCO Institute for Statistics and the Trends in International Migration published by the OECD.

GDP per capita and the Gini coefficient represent the income situation of the country. GDP per capita is an indicator for the general availability of resources and the Gini coefficient gives a picture of income inequality.

Moreover, immigrants in high-income countries with a lower degree of income inequality may be negatively self-selected with respect to their ability. Unfortunately, there is no time variation in the Gini coefficient in the available data. It is assumed that income inequality has not changed substantially within the analyzed time period and the Gini coefficient is taken as fixed for each country.

For a small number of country years, the Trends in International Migration provides information on regions where the migrants in a country come from. This information is used to account for the possibility that unobserved characteristics of the foreign population change over time.

Unfortunately, the used country data are incomplete, some missing values are imputed from other years, and unavailable data that cannot be imputed from other years are not dropped from the sample, but missing dummies are included in the regressions.

4 Results

This section contains the results of the baseline specification for all countries, the results for the OECD sample, and a number of robustness checks.

4.1 Baseline specification for all countries

Table 4 gives the estimation results in math and science for all countries. The first two columns of figures contain the results of the pooled WLS estimations, followed by the country-group effects and country-fixed-effects models. The regressions are weighted with the inverse standard error of the underlying decomposition and cluster robust (countries) standard errors are reported. The dependent variable is integration of immigrants and second-generation immigrants in math and science, respectively. The effects of income inequality, as well as entry age of schooling, cannot be estimated with country-fixed effects because these variables do not change over time.

Educational institutions Segregation of migrant students among schools and social segregation show no significant effects in the first four columns. All point estimates have negative signs. In mathematics, the coefficient of migrant segregation gets positive significant and the coefficient for social segregation remains negative and gets significant when differencing out country effects. In science, both estimates are not significant. The results suggest that migrant segregation is good and social segregation is bad for immigrant children, though both variables are significantly positively correlated. Migrant segregation gets insignificant and social segregation less significant if both coefficients are estimated separately. In total, the results are inconsistent and do not allow for any conclusion of wether segregation in schools is negatively or positively related to the integration of migrants. As mentioned above, TIMSS measures

Table 4 Results for all countries

Dependent variable: Integration	Pooled		Country-group-effects		Country-fixed-effects	
	Math	Science	Math	Science	Math	Science
Educational institutions						
Migrant segregation	−9.731	−15.210	−17.075	−16.995	54.727	26.979
	(9.918)	(12.335)	(13.441)	(11.953)	(31.336)[a]	(33.818)
Social segregation	−14.972	−32.242	−14.844	−8.452	−82.987	−55.405
	(18.224)	(20.839)	(30.081)	(29.827)	(40.715)[b]	(60.378)
Hours per year	16.961	11.709	19.554	15.176	16.171	11.973
	(5.754)[c]	(5.326)[b]	(5.943)[c]	(4.501)[c]	(7.787)[b]	(4.671)[b]
Hours per year sq	−0.879	−0.658	−0.972	−0.818	−0.739	−0.656
	(0.277)[c]	(0.293)[b]	(0.320)[c]	(0.242)[c]	(0.362)[b]	(0.221)[c]
Pre-prim enrollment	0.038	0.136	0.013	0.216	0.770	0.185
	(0.081)	(0.079)[a]	(0.101)	(0.076)[c]	(0.417)[a]	(0.260)
Age primary 5	11.145	8.334	6.918	8.680		
	(5.108)[b]	(3.407)[b]	(6.552)	(4.312)[b]		
Age primary 7	−4.045	−6.796	−4.013	2.331		
	(4.066)	(3.551)[a]	(6.522)	(5.126)		
Pupil–teacher ratio	−0.459	−0.603	0.178	0.281	0.397	1.030
	(0.252)[a]	(0.370)	(0.343)	(0.372)	(0.550)	(0.498)[b]
External studass	1.975	8.643	−6.176	−0.599	−6.768	13.725
	(6.690)	(6.465)	(7.444)	(6.113)	(23.931)	(29.356)
Remedial courses	30.353	16.481	28.604	20.694	15.258	−8.700
	(8.058)[c]	(8.617)[a]	(10.193)[c]	(8.105)[b]	(21.781)	(18.087)
Enrichment courses	−10.221	−6.912	−11.889	−4.565	−41.826	−13.393
	(5.859)[a]	(6.509)	(7.451)	(6.550)	(11.415)[c]	(16.650)
Income situation						
GDP	−0.265	−0.228	−0.154	−0.132	0.116	0.183
	(0.053)[c]	(0.056)[c]	(0.068)[b]	(0.065)[b]	(0.162)	(0.176)
Gini	−82.019	−30.598	−59.952	−54.084		
	(27.532)[c]	(33.354)	(32.825)[a]	(31.608)[a]		
GDP*Gini	0.643	0.377	0.398	0.248		
	(0.146)[c]	(0.161)[b]	(0.157)[b]	(0.168)		
Control variables						
National score	0.077	0.165	0.083	0.197	0.170	0.019
	(0.029)[c]	(0.028)[c]	(0.053)	(0.042)[c]	(0.198)	(0.200)
Second generation	6.079	9.642	5.528	9.420	5.294	9.083
	(3.177)[a]	(3.500)[c]	(3.457)	(3.704)[b]	(3.785)	(4.132)[b]
Wave dummies	Yes	Yes	Yes	Yes	Yes	Yes
Country-groups	No	No	Yes	Yes	No	No
C-groups*waves	No	No	Yes	Yes	Yes	Yes
Country dummies	No	No	No	No	Yes	Yes
Observations	334	334	334	334	334	334
R-squared	0.302	0.311	0.398	0.440	0.542	0.519
Adj R-squared	0.240	0.250	0.252	0.304	0.317	0.282

Weighted least squares, cluster robust standard errors in parentheses (clusters are countries), missing dummies included

[a] Statistical significance at 10%

[b] Statistical significance at 5%

[c] Statistical significance at 1%

segregation among classes and PISA refers to schools. Both segregation measures were interacted with TIMSS and PISA dummies, but no systematic differences have been found.

Time in school is represented by the mean number of instructional hours per school year (divided by 100). The variable is included in a quadratic form to allow for non-linear returns to schooling hours. The coefficients are statistically significant in all regressions and show the expected signs. According to the country-fixed effects model, time in school and integration are positively correlated up to about 1.094 h in math and 913 h in science. Given a mean value of 932 instructional hours per school year, a reform towards full-time schooling might be beneficial for migrant students in many countries.

Pre-primary school enrollment and integration are positively correlated; however, the effects are not always significant. The estimates of school entry-age are consistent with these results. Education systems where children start schooling at the age of five, compared to six, are associated with higher levels of integration. The magnitude of the coefficients is about ten test score points, capturing approximately 30% of the standard deviation of the dependent variable. Overall, the results suggest that early education, either in pre-primary or primary schools, is important for the educational assimilation of students with foreign background.

Mixed and mostly insignificant results are obtained for the pupil–teacher ratio and external student assessment. Furthermore, advancement activities for weak students show some significant positive effects, while enrichment courses show some significant negative coefficients. The availability of enrichment courses for gifted students was expected to be detrimental for integration if migrants are less likely to be accepted in such courses.[5]

Income situation   Both sets of regressions show that the income situation of the country is significantly correlated with the level of integration. This may be due to a resource effect, but also to the selection of economic migrants. The results indicate that high-income countries show a lower level of integration. Furthermore, a higher level of income inequality reduces integration of foreign students. This result was expected, since migrants often belong to the poor part of society and unequally spent resources should affect them negatively. Interestingly, the interactions of GDP and Gini show that foreign students are better integrated in high-income countries with a higher level of income inequality. This result may represent the selectivity of economic migrants. This is exactly what economic theory about migration predicts. Remember, economic migrants with high abilities are more likely to migrate to countries where they earn more. Thus, migrants are positively self-selected in high-income countries with a higher level of income inequality.

---

[5]I found heterogenous effects across country-groups for pre-primary enrollment and external student assessment. The coefficients of pre-primary education are statistically and economically most significant in Southern Europe and the English-speaking countries. Furthermore, the country-fixed-effects estimations show that external student assessment and integration in science are positively related in Southern Europe and the English-speaking countries, while negative associations are found in the poor Eastern European countries and the Near East. For the other variables, no systematic differences have been found.

Control variables The national average test score is significantly positively related to the integration of migrants in three out of six estimations. However, the sign is always positive. Interpreting the average test score as a quality measure of the education system, it appears that high-quality systems further migrant students to a greater extent. Moreover, second-generation immigrants do better than immigrants, both in math and in science. The gaps between immigrants of the first and second generation are of equal magnitudes in all regressions, about six test score points in mathematics and nine in science. Educational assimilation seems to be easier in math as compared to science. One explanation for this result is that math scores should be more related to IQ and science scores to factual knowledge; thus, they should be more influenced by language skills and social environment.

4.2 Baseline specification for the OECD sample

The identifying assumption of the model is more likely satisfied if one compares similar countries only. This is particularly important for the pooled and country-group effects specifications. Table 5 gives the results for the OECD countries.

Both segregation indices, again, show no consistent results. The estimates of hours per school year indicate the same pattern as above; however, the effects are statistically not significant in two out of six regressions. According to the country-fixed-effects specification, schooling time is positively associated with integration in math up to 1,244 instructional hours. The average within the OECD is about 954 h per year, implying considerable potentiality in a number of countries if the effect was causal. The magnitude of the effect is sizeable as well: starting from the mean, an increase in hours by 100 (approximately one standard deviation within the OECD) is associated with a higher level of integration of 8.5 test score points in math.

The coefficients of pre-primary education are statistically and economically more significant within the OECD and show an important magnitude in the fixed-effects math regression. According to this specification, a higher total enrollment rate by 25% points (one standard deviation within the OECD) is related to reinforced integration of about 34 math points, which is about one standard deviation of the dependent variable. Moreover, the coefficients on school entry at age five show the expected signs, but are not always statistically significant. The coefficients on age seven are mixed and not consistent.

While most of the other variables show mixed and insignificant results, external student assessment seems to be important within the OECD. Integration in science is higher if schools are not responsible for assessment policies. In the country-fixed-effects specification, a large coefficient is obtained that is statistically significant at the 1% level. If this coefficient would be causal, an increase in the fraction of schools that do not have the primary responsibility for student assessment policies by 0.2, this is one standard deviation in the OECD, could be associated with an increase in integration of about 12 science points.

Table 5 Results for OECD countries

Dependent variable: Integration	Pooled		Country-group-effects		Country-fixed-effects	
	Math	Science	Math	Science	Math	Science
Educational institutions						
Migrant segregation	−19.367	9.698	−15.036	9.542	43.685	84.580
	(19.984)	(16.818)	(19.855)	(18.894)	(44.388)	(45.065)[a]
Social segregation	−37.361	−64.743	13.725	33.631	−3.015	−22.148
	(23.264)	(31.485)[b]	(36.768)	(45.136)	(36.191)	(60.510)
Hours per year	36.105	23.887	38.093	20.390	44.272	9.898
	(13.240)[b]	(12.112)[a]	(14.501)[b]	(14.631)	(18.370)[b]	(13.649)
Hours per year sq	−1.879	−1.129	−1.789	−0.677	−1.779	−0.112
	(0.692)[b]	(0.660)[a]	(0.759)[b]	(0.773)	(0.951)[a]	(0.676)
Pre-prim enrollment	0.099	0.258	0.266	0.442	1.344	0.588
	(0.133)	(0.112)[b]	(0.141)[a]	(0.123)[c]	(0.517)[b]	(0.522)
Age primary 5	10.339	12.679	12.677	12.977		
	(5.812)[a]	(4.447)[c]	(8.080)	(9.041)		
Age primary 7	−8.918	0.258	2.309	18.666		
	(7.537)	(6.038)	(9.976)	(6.315)[c]		
Pupil–teacher ratio	0.052	−0.156	−0.032	0.341	0.048	1.239
	(0.414)	(0.556)	(0.543)	(0.737)	(0.699)	(1.033)
External studass	−1.390	17.342	9.079	33.882	3.930	59.793
	(10.103)	(9.307)[a]	(10.996)	(10.915)[c]	(27.129)	(19.316)[c]
Remedial courses	22.250	16.917	20.281	15.774	22.536	−23.939
	(15.241)	(13.386)	(18.210)	(12.873)	(32.590)	(24.955)
Enrichment courses	−8.118	−3.985	−13.505	−12.442	−80.203	−4.872
	(8.427)	(11.006)	(13.024)	(12.266)	(19.611)[c]	(18.268)
Income situation						
GDP	−0.361	−0.517	−0.500	−0.687	0.031	0.029
	(0.099)[c]	(0.124)[c]	(0.247)[a]	(0.292)[b]	(0.178)	(0.149)
Gini	−170.575	−204.598	−379.531	−533.591		
	(73.134)[b]	(83.645)[b]	(225.129)	(288.308)[a]		
GDP*Gini	0.928	1.129	1.438	1.769		
	(0.280)[c]	(0.340)[c]	(0.743)[a]	(0.891)[a]		
Control variables						
National score	−0.026	−0.059	0.156	0.200	0.249	0.370
	(0.107)	(0.097)	(0.193)	(0.108)[a]	(0.222)	(0.295)
Second generation	−2.321	3.151	−2.315	3.074	−2.052	2.924
	(3.059)	(3.694)	(3.215)	(3.886)	(3.435)	(4.216)
Wave dummies	Yes	Yes	Yes	Yes	Yes	Yes
Country-groups	No	No	Yes	Yes	No	No
C-groups*waves	No	No	Yes	Yes	Yes	Yes
Country dummies	No	No	No	No	Yes	Yes
Observations	190	190	190	190	190	190
R-squared	0.182	0.222	0.336	0.350	0.459	0.427
Adj R-squared	0.057	0.104	0.140	0.158	0.207	0.161

Weighted least squares, cluster robust standard errors in parentheses (clusters are countries), missing dummies included

[a]Statistical significance at 10%

[b]Statistical significance at 5%

[c]Statistical significance at 1%

The income situation gives the same picture as above. This result can be interpreted as a selection mechanism. Immigrants and second-generation immigrants in high-income countries with a higher degree of income inequality

are better off, which is consistent with the predictions of economic theory on the selectivity of economic migrants.

Overall, when estimating the model with OECD countries, most effects on institutions that have been obtained for all countries are corroborated and additional insights for the OECD countries are won.

4.3 Robustness checks

Four types of robustness checks are carried out. The model is estimated including controls for the composition of the migrant population with respect to the immigrants' home countries (see Table 6). Furthermore, the sensitivity of the results to the inclusion of language proficiency in the achievement regressions, to some sample restrictions in the micro-data and to the standardization of the unexplained differentials, are investigated. These estimations are not shown in the paper but results can be obtained from the author.

Controls for migration regions The first robustness check is based on a model that controls for the regions where the foreign population of a given country comes from. This strategy should remove the remaining problem that unobserved country attributes, like the composition of the immigrant population, change over time and may be correlated with changes in institutional characteristics of the education system. The information on migration regions is only available for a small number of country-years. As mentioned above, the other observations are not dropped but a missing dummy is included in the regressions. The results are given in Table 6 and are very similar to those of Table 4, where migration regions are not included.

National language in the achievement regressions As discussed in Section 2, the unexplained achievement gap is substantially larger when national language proficiency is not included in the achievement regressions of the decompositions. Whether the true level of integration is obtained by filtering out differences in language skills or not is a matter of opinion, therefore, I examined whether the results on institutional effects are sensitive to this decision. The results of schooling time are robust to this robustness check. Pre-primary education shows the same coefficients; however, the coefficient in the math country-fixed-effects regression is statistically not significant. Furthermore, school entry at age five and some coefficients on remedial and enrichment courses are not significant anymore.

Grade-age restrictions As mentioned in the Appendix, the grade and age variables in the achievement regressions might be endogenous if selection into grades differs between migrants and native students. To test for this bias, I restricted the sample to those grades with more then 30% of PISA students and furthermore dropped the oldest and youngest 5% of TIMSS students. Some

Table 6 Results with migration regions

Dependent variable: Integration	Pooled		Country-group-effects		Country-fixed-effects	
	Math	Science	Math	Science	Math	Science
Educational institutions						
Migrant segregation	−11.416	−19.994	−5.566	−8.483	49.431	11.384
	(11.237)	(13.725)	(18.082)	(15.212)	(40.444)	(40.969)
Social segregation	−6.901	−29.013	−19.984	−19.199	−65.691	−2.383
	(19.548)	(19.731)	(31.063)	(31.089)	(53.554)	(69.557)
Hours per year	16.754	11.710	19.519	15.327	17.261	9.860
	(6.049)[a]	(5.683)[b]	(5.745)[a]	(4.789)[a]	(8.251)[b]	(5.207)[c]
Hours per year sq	−0.848	−0.638	−0.951	−0.812	−0.847	−0.604
	(0.291)[a]	(0.317)[b]	(0.308)[a]	(0.254)[a]	(0.406)[b]	(0.284)[b]
Pre-prim enrollment	0.037	0.134	−0.056	0.172	0.734	0.016
	(0.096)	(0.089)	(0.125)	(0.088)[c]	(0.367)[b]	(0.259)
Age primary 5	7.858	15.422	13.945	22.507		
	(7.136)	(5.739)[a]	(10.114)	(7.516)[a]		
Age primary 7	−3.955	−6.395	−7.262	0.029		
	(4.684)	(4.004)	(6.666)	(5.752)		
Pupil–teacher ratio	−0.404	−0.537	0.280	0.337	0.324	1.092
	(0.324)	(0.433)	(0.443)	(0.409)	(0.669)	(0.600)[c]
External studass	3.869	13.767	−6.491	3.678	1.615	3.950
	(8.337)	(8.232)[c]	(9.801)	(8.174)	(29.047)	(37.861)
Remedial courses	30.125	13.253	24.039	13.332	22.182	1.525
	(8.776)[a]	(9.813)	(11.460)[b]	(10.526)	(23.642)	(20.337)
Enrichment courses	−12.833	−6.216	−12.007	−3.719	−41.000	−14.409
	(6.497)[c]	(7.066)	(7.780)	(7.120)	(13.217)[a]	(15.613)
Income situation						
GDP	−0.262	−0.195	−0.140	−0.060	0.170	0.332
	(0.065)[a]	(0.057)[a]	(0.083)[c]	(0.073)	(0.241)	(0.238)
Gini	−89.593	−25.918	−46.700	−36.371		
	(29.362)[a]	(33.920)	(32.229)	(30.470)		
GDP*Gini	0.642	0.331	0.398	0.121		
	(0.189)[a]	(0.182)[c]	(0.194)[b]	(0.202)		
Control variables						
National score	0.078	0.162	0.094	0.194	0.180	0.110
	(0.033)[b]	(0.032)[a]	(0.054)[c]	(0.044)[a]	(0.188)	(0.189)
Second generation	6.085	9.811	5.485	9.326	5.333	8.934
	(3.239)[c]	(3.547)[a]	(3.539)	(3.791)[b]	(3.865)	(4.206)[b]
Wave dummies	yes	yes	yes	yes	yes	yes
Country-groups	no	no	yes	yes	no	no
C-groups*waves	no	no	yes	yes	yes	yes
Country dummies	no	no	no	no	yes	yes
F-Statistics (migration regions)	2.05	1.82	1.34	1.75	0.88	1.44
Observations	334	334	334	334	334	334
R-squared	0.309	0.319	0.410	0.449	0.550	0.530
Adj R-squared	0.225	0.236	0.241	0.291	0.300	0.268

Weighted least squares, cluster robust standard errors in parentheses (clusters are countries), missing dummies included and migration regions included

[a] Statistical significance at 1%

[b] Statistical significance at 5%

[c] Statistical significance at 10%

coefficients loose statistical significance; however, all significant effects show the same signs and magnitudes as in the baseline model.

Robustness to the standardization Since the unexplained differential is calculated by using the average endowments of foreign students in each country, a standardization was needed to reach comparability. I used the mean characteristics of foreign students in the whole sample. To check whether the results are influenced by this choice, especially the results for the OECD sample, I replaced the standardization vector by the average characteristics of migrants within the OECD. The results are very robust to this sensitivity check.

5 Conclusions

This essay is aimed at quantifying the drawback of students with migration background in secondary schools. Blinder–Oaxaca decompositions show that, on average, the test score gaps between students with foreign background and natives cannot be entirely explained with differences in the students' productivity characteristics. In most countries, a part of the test score gap remains unexplained. As shown in Figs. 1 and 2, educational gaps between native and foreign students and the shares that remain unexplained vary substantially across country groups and OECD countries.

The English-speaking countries or traditional countries of immigration are found in the middle and lower parts of the gap distribution, and here, the unexplained differential is almost always positive, meaning that immigrant children get higher returns to their characteristics than natives. The achievement gaps in Middle and Northern Europe are largest and average out to -33 math and -35 science points, with 15% and 26% remaining unexplained.

The proficiency of the national language turned out to have significant effects on integration. The unexplained test score gap is substantially larger if national language is not included in the achievement regressions. This finding is not new and strongly argues for public policies that encourage immigrants to learn the national language.

In the second part, integration of students with migration background is related to institutional characteristics of the education system, the income situation of the country, and some control variables. Interestingly, the estimated coefficients on the income situation show exactly the results predicted by economic theory of migration: high-income countries with a high level of income inequality should attract immigrants with higher abilities. In fact, educational integration is higher in high-income countries with higher levels of income inequality.

Educational institutions were found to be correlated with the integration of migrants. Specifically, time in school and early education are positively related to integration in both math and science. The result on pre-primary education is in line with the studies on Head Start and indicate that promotion of students with migration background should start as early as possible.

Central examinations, measured by the fraction of schools that do not have the primary responsibility for student assessment policies, are associated with higher academic science achievement of migrant students in the OECD. This is consistent with other studies and may be due to the involved restriction in the latitude of teachers' grading practices and information-induced pressure on students, teachers, and schools.

Overall, the results of the study suggest that the design of the education system is important for children with migration background. Recent demographic trends in many industrialized countries indicate that the integration of migrants will be a major challenge in future. To meet this challenge, policy makers should regard educational integration as an important precondition and education policy as a main instrument to further the economic assimilation of immigrants in the host societies.

Appendix

A1 Comparability of PISA and TIMSS

Although the surveys are very similar, some aspects differ between PISA and TIMSS. TIMSS measures the mastery of an internationally agreed curriculum, while PISA focuses on challenges of every-day life. The test questions span the same topics but differ in their reference to reality. Furthermore, the target population consists of teenagers in secondary education. TIMSS covers children in the grade(s) with the highest proportion of 13-year-olds (typically grade 7 or 8) and PISA covers 15-year-olds, independent of grade. It might be the case that immigrants are more likely placed in lower grades, given their age, for two reasons: they might have started later with school or repeated a grade. The inclusion of age and grade in the educational production functions should mitigate this problem. Furthermore, a sensitivity check focusing on this issue is presented in Section 4.3.

A2 Student achievement scores

The test scores in PISA and TIMSS were standardized to a mean of 500 and a standard deviation of 100, but based on a different set of countries. While PISA focuses on OECD countries, in TIMSS, participating countries are more heterogenous. Thus, a typical OECD country is likely to perform above average in TIMSS, but not in PISA, and the test scores cannot be compared without transformation.

Fifteen countries participated in both PISA 2003 and TIMSS 2003. Given the similar survey designs, it is assumed that the test score distribution in TIMSS should be equal to that in PISA in these 15 countries. Thus, the PISA scores of the common subsample were transformed to the same distribution as the TIMSS subsample. The TIMSS achievement scores are directly comparable across waves; thus, no further step for TIMSS was necessary. Because of the alternate major subject assessment in each wave, the PISA scores are directly comparable across waves in science, but not in mathematics. Thus, the science scores where just added to the scale, and for math, a second step was necessary. I calculated the score distribution of the PISA 2003 data for the subsample of countries participating in 2000 and 2003 and applied this distribution to the PISA 2000 subsample.

After this transformation procedure, the math and science scales were transformed to a mean and standard deviation of 500 and 100. Note that the transformation has no influence on the ranking of the students and does not change the distance in terms of standard deviations between any two students. Very similar approaches were used by Hanushek and Wößmann (2005), Schütz et al. (2008), and Ammermüller (2005).

References

Aksoy T, Link CR (2000) A panel analysis of student mathematics achievement in the US in the 1990s: does increasing the amount of time in learning activities affect math achievement? Econ Educ Rev 19(3):261–277

Ammermüller A (2005) Educational opportunities and the role of institutions. Discussion Paper No 05–44, Center for European Economic Research, Mannheim

Ammermüller A (2007) Poor background or low returns? Why immigrant students in Germany perform so poorly in PISA. Educ Econ 15(2):215–230

Angrist JD, Krueger AB (1992) The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples. J Am Stat Assoc 87(418):328–336

Angrist JD, Lavy V (1999) Using Maimonides' rule to estimate the effect of class size on scholastic achievement. Q J Econ 114(2):533–575

Blinder AS (1973) Wage discrimination: reduced form and structural estimates. J Hum Resour 8(4):436–455

Borjas GJ (1987) Self-selection and the earnings of immigrants. Am Econ Rev 77(4):531–553

Borjas GJ (1999) Economic analysis of immigration. In: Ashenfelter O, Card D (eds) Handbook of labor economics, vol 3A. Elsevier, Amsterdam, pp 1697–1761

Brown G, Micklewright J, Schnepf SV, Waldmann R (2007) Cross-national surveys of learning achievement: how robust are the findings? J R Stat Soc Ser A 170(3):623–646

Burgess S, Wilson D (2005) Ethnic segregation in England's schools. Trans Inst Br Geogr 30(1): 20–36

Card D, Krueger AB (1992) Does school quality matter? Returns to education and the characteristics of public schools in the United States. J Polit Econ 100(1):1–40

Carneiro P, Heckman JJ, Masterov DV (2005) Labor market discrimination and racial differences in premarket factors. J Law Econ 48(1):1–39

Chiswick BR (1999) Are immigrants favorably self-selected? American Economic Review Papers and Proceedings 89(2):181–185

Chiswick BR, Miller PW (2008) Why is the payoff to schooling smaller for immigrants? Labour Econ 15(6):1317–1340

Currie J (2001) Early childhood education programs. J Econ Perspect 15(2):213–238

Currie J, Thomas D (1999) Does head start help hispanic children? J Public Econ 74(2):235–262

Duncan OD, Duncan B (1955) A methodological analysis of segregation indexes. Am Sociol Rev 20(2):210–217

Entorf H, Minoiu N (2005) What a difference immigration policy makes: a comparison of PISA scores in Europe and traditional countries of immigration. Ger Econ Rev 6(3):355–376

Epple D, Newlon E, Romano R (2002) Ability tracking, school competition and the distribution of educational benefits. J Public Econ 83(1):1–48

Hanushek EA, Luque JA (2003) Efficiency and equity in schools around the world. Econ Educ Rev 22:481–502

Hanushek EA, Wößmann L (2005) Does educational tracking affect performance and inequality? Difference-in-difference evidence across countries. Econ J 116(510):C63–C76

Hartog J, Zorlu A (2009) How important is homeland education for refugees' economic position in the Netherlands. J Popul Econ 22(1):219–246

Jenkins SP, Micklewright J, Schnepf SV (2008) Social segregation in secondary schools: how does England compare with other countries? Oxf Rev Educ 34(1):21–37

Krueger AB (1999) Experimental estimates of education production functions. Q J Econ 114(2):497–532

Lewis KA, Seidman LS (1994) Math-time capital matters: a cross-country analysis. Econ Educ Rev 13(3):215–226

Miller PW (1999) Immigration policy and immigrant quality: the Australian points system. American Economic Review Papers and Proceedings 89(2):192–197

Oaxaca R (1973) Male-female wage differentials in urban labor markets. Int Econ Rev 14(3):693–709

OECD (2001) Knowledge and skills for life: first results from PISA 2000. Organisation for Economic Co-operation and Development, Paris

OECD (2002) PISA 2000. Technical report, Organisation for Economic Co-operation and Development, Paris

OECD (2004) Learning for tomorrow's world: first results from PISA 2003. Organisation for Economic Co-operation and Development, Paris

OECD (2005) PISA 2003. Technical report, Organisation for Economic Co-operation and Development, Paris

Rees DI, Argys LM, Brewer DJ (1996) Tracking in the United States: descriptive statistics from NELS. Econ Educ Rev 15(1):83–89

Sacerdote B (2001) Peer effects with random assignment: results for Dartmouth roommates. Q J Econ 116(2):681–704

Schneeweis N, Winter-Ebmer R (2007) Peer effects in Austrian schools. Empir Econ 32(2–3):387–409

Schütz G, Ursprung HW, Wößmann L (2008) Education policy and equality of opportunity. Kyklos 61(2):279–308

Silber J, Weber M (1999) Labor market discrimination: are there significant differences between the various decomposition procedures? Appl Econ 31(3):359–365

Winston GC, Zimmerman DJ (2003) Peer effects in higher education. NBER Working Paper Nr. 9501

Wößmann L (2005a) Educational production in Europe. Econ Policy 20(43):445–504

Wößmann L (2005b) The effect heterogeneity of central examinations: evidence from TIMSS, TIMSS-Repeat and PISA. Educ Econ 13(2):143–169